Stuart J. Cooper Co. Recycled Paper
213-747-7141 · 800-421-8703 · Fax 213-747-3035

Ex-8

Cell, Vol. 37, 767–778, July 1984, Copyright © 1984 by MIT

0092-8674/84/070767-12 $02.00/0

The Structure of an Antigenic Determinant in a Protein

Ian A. Wilson, Henry L. Niman, Richard A. Houghten, Andrew R. Cherenson, Michael L. Connolly and Richard A. Lerner
Department of Molecular Biology
Research Institute of Scripps Clinic
La Jolla, California, 92037

Summary

The immunogenic and antigenic determinants of a synthetic peptide and the corresponding antigenic determinants in the parent protein have been elucidated. Four determinants have been defined by reactivity of a large panel of antipeptide monoclonal antibodies with short, overlapping peptides (7–28 amino acids), the immunizing peptide (36 amino acids), and the intact parent protein (the influenza virus hemagglutinin, HA). The majority of the antipeptide antibodies that also react strongly with the intact protein recognize one specific nine amino acid sequence. This immunodominant peptide determinant is located in the subunit interface in the HA trimeric structure. The relative inaccessibility of this site implies that antibody binding to the protein is to a more unfolded HA conformation. This antigenic determinant differs from those previously described for the hemagglutinin and clearly demonstrates the ability of synthetic peptides to generate antibodies that interact with regions of the protein not immunogenic or generally accessible when the protein is the immunogen.

Introduction

Despite recent advances in molecular immunology, the detailed chemistry of antigen–antibody union and a complete description of the exact structure of antigenic determinants of proteins remained undetermined. Most of the early studies of antigenic determinants followed similar experimental protocols in which the reactivity of antibodies made to intact protein structures was determined on denatured and/or fragmented molecules (reviewed by Crumpton, 1974; Benjamin et al., 1984). These methods led to proposals for the complete antigenic structures of myoglobin (Atassi, 1975), hen egg white lysozyme (Atassi, 1978), and a nearly complete antigenic structure of serum albumin (Atassi et al., 1979). These antigenic sites were identified as either continuous, comprised of short linear sequences, as in myoglobin, or, more frequently, discontinuous, comprised of residues separated in the amino acid sequence, as in lysozyme. However, more recent work with antisera to specific antigenic sites has challenged the identification of these immunochemically defined antigenic structures (reviewed by Benjamin et al., 1984). For example, Hurrel et al. (1978) reported that goat antisera to sperm whale myoglobin were not at all inhibited by beef myoglobin which should share the same determinant, residues 56–62. Also, goat antisera had vastly different affinities for pig and horse myoglobins with again, identical sequences in previously reported antigenic sites, 56–62 and 94–99. Similar experiments with myoglobin have been reported by East et al. (1980).

Recently monoclonal antibodies have been used to map the antigenic areas of hen egg white lysozyme (Smith-Gill et al., 1982) and sperm whale myoglobin (Berzofsky et al., 1982). For example, the major lysozyme antigenic site (epitope) was proposed to include a region with dimensions of at least 13 × 6 × 15Å which was shared by lysozymes of seven related galliform species and was also outside the previously suggested antigenic sites (Atassi, 1978).

A combination of three-dimensional structure analysis and virus neutralization studies has also been used to elucidate the immunogenic areas of a viral protein. X-ray structural determination of influenza virus hemagglutinin (A/Aichi/2/68, H3 subtype) by Wilson et al. (1981) led to a proposal of four antibody-combining sites which were implicated in viral neutralization (Wiley et al., 1981). Variation in amino acid sequence of field mutants from 1968 to 1979 (reviewed by Laver et al., 1980; Both et al., 1983) and of variants selected in the presence of monoclonal antibodies (Laver et al., 1979; Webster and Laver, 1980; Wiley et al., 1984) led to proposals for the amino acids involved in antibody binding and estimations of the approximate dimensions for two of the sites (Wiley et al., 1981, 1984). The antigenic structure of a different influenza virus subtype (H1, A/Puerto Rico/8/34) hemagglutinin was determined by mapping with monoclonal antibodies (Caton et al., 1982) and the antigenic sites were broadly consistent with those reported for the H3 subtype (Wiley et al., 1981, 1984; Underwood, 1982; Daniels et al., 1984).

All the above approaches depend on defining antigenic determinants after immunization with the whole protein. Consequently, it is difficult to localize precisely the antibody-combining site and to define the extent and number of amino acids directly involved in antibody–antigen union. Changes in amino acids that alter antibody binding and specificity may not always be easy to interpret in the context of complex structures. For example, changes in one region of the protein may in some circumstances alter the conformation of quite distant regions.

The generation of antipeptide antibodies that have predetermined sequence specificity would seem to offer an alternative approach to the problem of understanding the general nature and structure of antigenic determinants and antibody–antigen interaction. Extensive use of synthetic peptides has demonstrated their ability to induce antibodies with predetermined sequence specificity that bind to almost any region of a protein (e.g., Green et al., 1982) and recent studies have shown that a large percentage of monoclonal antibodies raised against synthetic peptides cross-react with the intact protein molecule (Niman et al., 1983). Such antibodies can define antigenic determinants in proteins in that the antibodies bind to the intact protein as well as to the synthetic peptide. Since the antibodies

are not generated against the native protein, we refer to the determinants as *in* a protein rather than *of* a protein. The study of antigenic determinants in a protein defined by antipeptide antibodies differs from the study of immunogenic determinants of proteins defined by immunization with the intact protein in that, as mentioned above, the latter determinants are often more complex and frequently composed of amino acids from distant parts of the polypeptide chain. This complexity makes these determinants less amenable to structural study.

In the present report, the structure of antigenic determinants in a protein has been determined by using monoclonal antibodies that bind to short synthetic peptides and their cognate sequences in the intact folded protein. Since some of the antibody binding sites are essentially inaccessible in the native trimeric molecule, our data suggest that conformational changes in the protein may precede and/or accompany antigen–antibody union. This could imply either that a conformational change is induced by antibody binding or that the antibody binds to a minor conformational state of the protein.

Illustrations of the location of these peptides in the A/Aichi/2/68 (X:31) hemagglutinin structure determined by Wilson et al. (1981) are presented. A new graphics program (Connolly, unpublished) is used to examine the location of these peptide sequences in the three-dimensional structure derived from the x-ray coordinates (Wilson, Skehel, and Wiley, unpublished data). This program extends the color raster display of solvent-accessible surfaces (Connolly, 1983a, 1983b). The new representation allows the surfaces to be interpreted easily by viewing a ball-and-stick model through its translucent molecular surface. The chemical structure of the protein can then be viewed simultaneously with its accessible surface.

Results

Characteristics of Antipeptide Antibodies Selected for Peptide Binding Studies

Monoclonal antibodies with predetermined sequence specificity were generated by immunizing mice with a synthetic peptide coupled to the carrier protein keyhole limpet hemocyanin. The synthetic peptide represents residues 75–110 in the HA1 chain of the H3 subtype X:47 (A/Victoria/3/75) hemagglutinin (Min Jou et al., 1980) when aligned with the H3 subtype X:31 (A/Aichi/2/68) sequence (Verhoeyen et al., 1980). Twenty-one different monoclonal antibodies were raised as described in Niman et al. (1983). The majority of these antibodies (16 out of 21) cross-reacted strongly with X:47 influenza virus in enzyme-linked immunosorbent assays (ELISA), immunoprecipitation, or Western blots (Table 1). Five antibodies did not cross react in either ELISA or immunoprecipitation, although two of these five did show cross reactivity under the more denaturing Western blot conditions (Niman et al., 1983). Thus most of these 21 antibodies recognized not only the peptide against which they were raised, but also the whole virus. The 18 antibodies were used to test further cross reactivity with the entire bromelain-released hemaggutinin and with a variety of smaller peptides to try to identify a more specific recognition site.

The Antipeptide Antibodies Cross-React with the Intact Hemagglutinin

The majority of the antipeptide antibodies cross-reacted strongly with X:47 bromelain-released hemagglutinin (BHA) in ELISA. The titers for the binding of antibodies to intact protein were in the same range as those for binding to the intact peptide (Table 1). Binding activity is expressed as amount of antigen, serially diluted and fixed on microtiter plates, which gives half-maximal binding when a constant amount (~100 fmoles) of antibody is added. The binding of the different monoclonals to BHA ranged from 50 fmoles to >6000 fmoles and binding to the 36 amino acid parent synthetic peptide varied from 60 to 1700 fmoles. Thus the monoclonals varied by two logs in their binding profiles. The antibodies that had weaker half-maximal binding also had lower maximal plateaus when recorded in ELISA. These results confirmed earlier identification of different reactivities of the antibodies with peptides and the virus by ELISA, immunoprecipitation, and Western blots (Niman et al., 1983). Hybridomas H23F10, H26D02 and H17B11 cross-reacted with virus in ELISA or Western blots, but not in immunoprecipitation assays. Hybridomas H17D12, H22F05, H17D09, and H16C08 had no detectable binding in any of these assays. The latter six hybridomas (i. e., excluding H23F10) had substantially lower maximal plateaus and weaker half-maximal binding in ELISA. Some of the antibodies did appear to bind better to the intact protein than the free peptide (e. g., H19D03, Table 1), although it is difficult to quantitate accurately the amount of each antigen bound to the microtiter plates.

Chemical Identification of Antibody-Binding Sites

The precise chemical identification of the antibody-binding sites is described in detail elsewhere (Niman et al., submitted). The synthetic peptides that specify the binding sites (1–4) are shown in Figure 1. The binding activities of the hybridomas are listed in Table 1. Analysis of the binding studies indicated that most of the monoclonal antibodies bind to a peptide with a sequence substantially shorter than the parent peptide. Several smaller, overlapping synthetic peptides further localized the sequences to which antibodies bound and showed binding titers approximately equivalent to that of the parent peptide.

Eight of the 18 monoclonals reacted with peptides confined to one short sequence of the parent molecule (Table 1, H23D02 through H26D08). The data indicate that these hybridomas bind to residues in the 36 amino acid synthetic peptide that correspond to HA1 residues 98–106. This sequence (site 4) is identified not by binding to a single peptide with sequence 98–106, but by binding to a series of overlapping peptides (Table 1, Figure 1). The use of overlapping segments is necessary because peptides of limited sequence length may not have all residues available for antibody binding when they are attached to

Table 1

Ab	Site	Virus	HA 1–503	1 75–110	2 52–86Y	3 Y75–82	4 75–84	5 75–86	6 Y75–89	7 79–87	8 83–92
23D02	4	+	100	70	>30000	>250000	>200000	>200000	>150000	>250000	>250000
21C08	4	+	140	70	>30000	>250000	>200000	>200000	>150000	>250000	>250000
19B10	4	+	490	140	>30000	>250000	>200000	>200000	>150000	>250000	>250000
20C01	4	+	310	160	>30000	>250000	>200000	>200000	>150000	>250000	>250000
24E07	4	+	180	90	>30000	>250000	>200000	>200000	>150000	>250000	>250000
16E07	4	+	190	160	>30000	>250000	>200000	>200000	>150000	>250000	>250000
26A09	4	+	50	60	>30000	>250000	>200000	>200000	>150000	>250000	>250000
26D08	4	+	120	80	>30000	>250000	>200000	>200000	>150000	>250000	>250000
20F04		+	390	1700	>30000	>250000	>200000	>200000	>150000	>250000	>250000
17D09		–	420	420	>30000	>250000	>200000	>200000	>150000	>250000	100000
17B11		0	2000	1300	>30000	>250000	>200000	>200000	>150000	>250000	>250000
16C08		–	>5900	1000	>30000	>250000	>200000	>200000	>150000	>250000	>250000
23F10	2	0	380	1400	900	>250000	>200000	20000	6500	3900	>250000
23B05	3	+	400	1600	>32000	>250000	>200000	>200000	>150000	>250000	3800
19D03	1	+	120	840	1100	>250000	>200000	6300	5900	>250000	>250000
26D02	4'	0	780	1300	>30000	>250000	>200000	>200000	>150000	>250000	>250000
22F05	4'	–	1500	250	>30000	>250000	>200000	>200000	>150000	>250000	>250000
17D12	4'	–	420	1700	>30000	>250000	>200000	>200000	>150000	>250000	>250000

Ab	Site	9 83–110	10 87–94	11 88–110	12 91–99	13 93–110	14 95–102	15 97–106	16 98–110
23D02	4	330	>550000	170	>250000	360	>300000	330	440
21C08	4	180	>550000	150	>250000	320	>300000	990	580
19B10	4	330	>550000	180	>250000	500	>300000	1300	370
20C01	4	460	>550000	200	>250000	630	>300000	1100	440
24E07	4	290	>550000	170	>250000	420	>300000	650	290
16E07	4	290	>550000	200	>250000	560	>300000	830	440
26A09	4	130	>550000	120	>250000	230	>300000	210	160
26D08	4	180	>550000	150	>250000	420	>300000	760	180
20F04		50000	>550000	>70000	>250000	>150000	>300000	>250000	>150000
17D09		>75000	>550000	>70000	>250000	>150000	>300000	>250000	>150000
17B11		>75000	>550000	>70000	>250000	>150000	>300000	>250000	>150000
16C08		>75000	>550000	>70000	>250000	>150000	>300000	>250000	>150000
23F10	2	>75000	>550000	>70000	>250000	>150000	>300000	>250000	>150000
23B05	3	7300	>550000	>70000	>250000	>150000	>300000	>250000	>150000
19D03	1	>75000	>550000	>70000	>250000	>150000	>300000	>250000	>150000
26D02	4'	15000	>550000	15000	>250000	65000	>300000	100000	150000
22F05	4'	15000	>550000	15000	>250000	65000	>300000	100000	150000
17D12	4'	9100	>550000	6900	>250000	65000	>300000	100000	150000

Ab	Site	17 98–104	18 98–105	19 98–106	20 99–106	21 100–106	22 100–108	23 103–110	24 104–139
23D02	4	400000	70000	3900	750	3200	2000	>300000	>65000
21C08	4	200000	60000	3900	790	2800	2000	>300000	>65000
19B10	4	200000	>350000	30000	300000	150000	4000	>300000	>65000
20C01	4	>400000	200000	30000	2800	4900	6700	>300000	>65000
24E07	4	200000	200000	3900	950	6500	2000	>300000	>65000
16E07	4	200000	200000	15000	2200	9700	4000	>300000	>65000
26A09	4	80000	20000	3900	260	2400	1700	>300000	>65000
26D08	4	100000	45000	7900	360	4900	2000	>300000	>65000
20F04		>400000	>350000	>250000	>300000	>400000	>200000	>300000	>65000
17D09		>400000	>350000	>250000	>300000	>400000	>200000	>300000	>65000
17B11		>400000	>350000	>250000	>300000	>400000	>200000	>300000	>65000
16C08		>400000	>350000	>250000	>300000	>400000	>200000	>300000	>65000
23F10	2	>400000	>350000	>250000	>300000	>400000	>200000	>300000	>65000
23B05	3	>400000	>350000	>250000	>300000	>400000	>200000	>300000	>65000
19D03	1	>400000	>350000	>250000	>300000	>400000	>200000	>300000	>65000
26D02	4'	400000	350000	250000	300000	400000	200000	>300000	>65000
22F05	4'	400000	350000	250000	300000	400000	200000	>300000	>65000
17D12	4'	400000	350000	250000	300000	400000	200000	>300000	>65000

Binding of antibody to peptides and BHA in ELISA assay. Values are presented as the amount of antigen in fmoles, serially diluted and fixed on microtiter plates, for half-maximal binding when a constant amount of antibody (~100 fmoles) is added. Some of the peptide antigens can dilute out equivalently such that they have apparently similar half-maximal binding quantities. However, the absolute maximal binding and plateau levels may differ in ELISA and are not represented by binding data represented only by serial dilution. Such lowered binding activities can be explained either by antibodies binding strongly to the peptide on the plate and the complex detaching from the microtiter plates or by weaker and more labile antibody binding to the antigen. For example, peptides 20 and 21 have these lower binding plateau profiles. Blocks of binding data used to designate the four sequences that specify the antigenic determinants (sites 1–4) are highlighted by bold vertical lines. Antibody binding to virus is shown for ELISA, Western blot, and immunoprecipitation assays. Significant binding to the virus in all three assay is indicated by a (+), binding in one or two of the assays by a (0), and no binding in any test by a (–). The peptide numbers are as shown in Figure 1 and the antibodies and virus titers are as described in Niman et al., 1983.

microtiter plates. Hence nested fragments around the binding site can specify the amino acids involved in antibody–antigen binding. Three hybridomas (H26D02, H22F05, and H17D12) also bind in this region but cannot be localized further than to the sequence 88–110.

Three different antibodies can be localized in their binding to sequences at the amino end of the parent peptide. These binding sites overlap and are identified as site 1, site 4. The remaining four monoclonals (H20F04, H17D09, H17B11, and H16808) have negligible or extremely low binding to any peptide shorter than the parent peptide and cannot be localized to any particular subsequence.

A Model for the Structure of the Antigenic Determinants of the Synthetic Peptide

Eleven out of 18 of the monoclonal antibodies recognize four short sequential segments contained in the

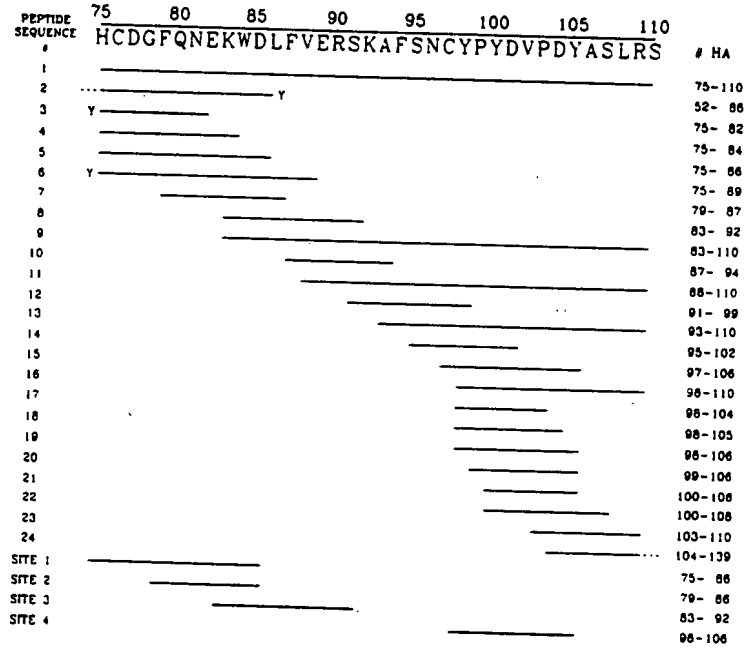

Figure 1. The Amino Acid Sequences of the Peptides Synthesized for Binding Studies to Anti-Peptide Monoclonal Antibodies

The parent sequences against which the monoclonal antibodies were raised correspond to residues 75–110 (X:31 numbering) of influenza virus X:47 HA1. Three peptides (2, 3, 6) have a tyrosine prepended or appended to the fragment (designated with a Y). Two peptides (2, 24) extend before or beyond the sequence 75–110. The sequences to which the majority of the antibodies are identified to bind in ELISA are designated sites (1–4).

one of two main areas of the parent peptide. Three sequences are around the amino terminus of the peptide (sites 1–3, residues 75–92) and one is near its carboxyl end (site 4, 98–106, Figure 1, Table 1). The data for these 11 monoclonals are consistent with a model in which the antigenic determinants consist of short, linear segments of polypeptide chain, spatially separated from the other residues in the peptide. It is interesting to note that the conformation that these segments adopt in the native molecule has sites (2–4) each consisting of a single stretch of extended chain sandwiched between bends (Figure 2). Site 1 varies somewhat from this structure in that it contains two extended chains connected by a bend at residue 80 (Figure 2). Four of the antibodies cannot be localized in binding studies to a short peptide and may recognize a more complex tertiary fold of the 36 amino acid peptide.

The immunodominant region, site 4, which corresponds to sequence 98–106 is shown on a model of the peptide, 75–110, in Figure 3. This antigenic determinant corresponds to no more than 25% of the total amino acids in the sequence 75–110 and in the native conformation has an accessible surface area of 763 sq. Å (24% of the total). Similarly, site 2, residues 79–86, has an accessible surface area of 800 sq. Å when calculated in the conformation it adopts in the native protein. The dimensions of sites 2 and 4 are approximately 13 × 22 × 12Å and 12 × 22 × 14Å, respectively.

These structural renderings represent an idealized situation for envisaging recognition by antibody in which both the free peptide and the intact protein can obtain the same conformation. The determination of the correctness of this or any other model must await further structural studies of the peptide and peptide–antibody complex. Whichever model turns out to be correct, it is already clear that the majority of these antipeptide antibodies recognize determinants which are short, linear segments of the polypeptide chain and differ markedly from the assembled topographical determinants generally described for the antigenic sites of proteins when the intact protein is the immunogen (Crumpton, 1974; Benjamin et al., 1984). In contrast, the four antibodies that cannot be localized in binding studies to a short peptide segment may recognize determinants which more closely resemble complex assembled topographical sites.

Location of the Synthetic Peptide in the Protein

Analysis of the structure of the antigenic determinants of the synthetic peptide is at present, as stated above, hampered by uncertainty abut the conformation of the free peptide in solution. However, many of these difficulties are abrogated by analyzing the structure of the peptide and the sequences defining the antigenic determinants in the native protein. The antibodies have approximately equivalent binding titers with the free peptide and intact protein (Table 1) and hence the antigenic determinants are shared by both.

The location in the hemaggutinin molecule corresponding to the 36 amino acid chemically synthesized peptide is central in the globular head region of the HA1 polypeptide chain (Figure 3, Wilson et al., 1981). This peptide consists of six segments of extended polypeptide chain connected by a series of bends. The amino and carboxyl ends of the chain contain three bends and a helix. The peptide contains residues which are either on the surface (75–83, 91–96), buried in the monomer (83–91, 97–99, 108–110), or essentially inaccessible in the trimer interface (100–107) (Figure 4). When considered as part of the native monomeric protein, the peptide has an accessible

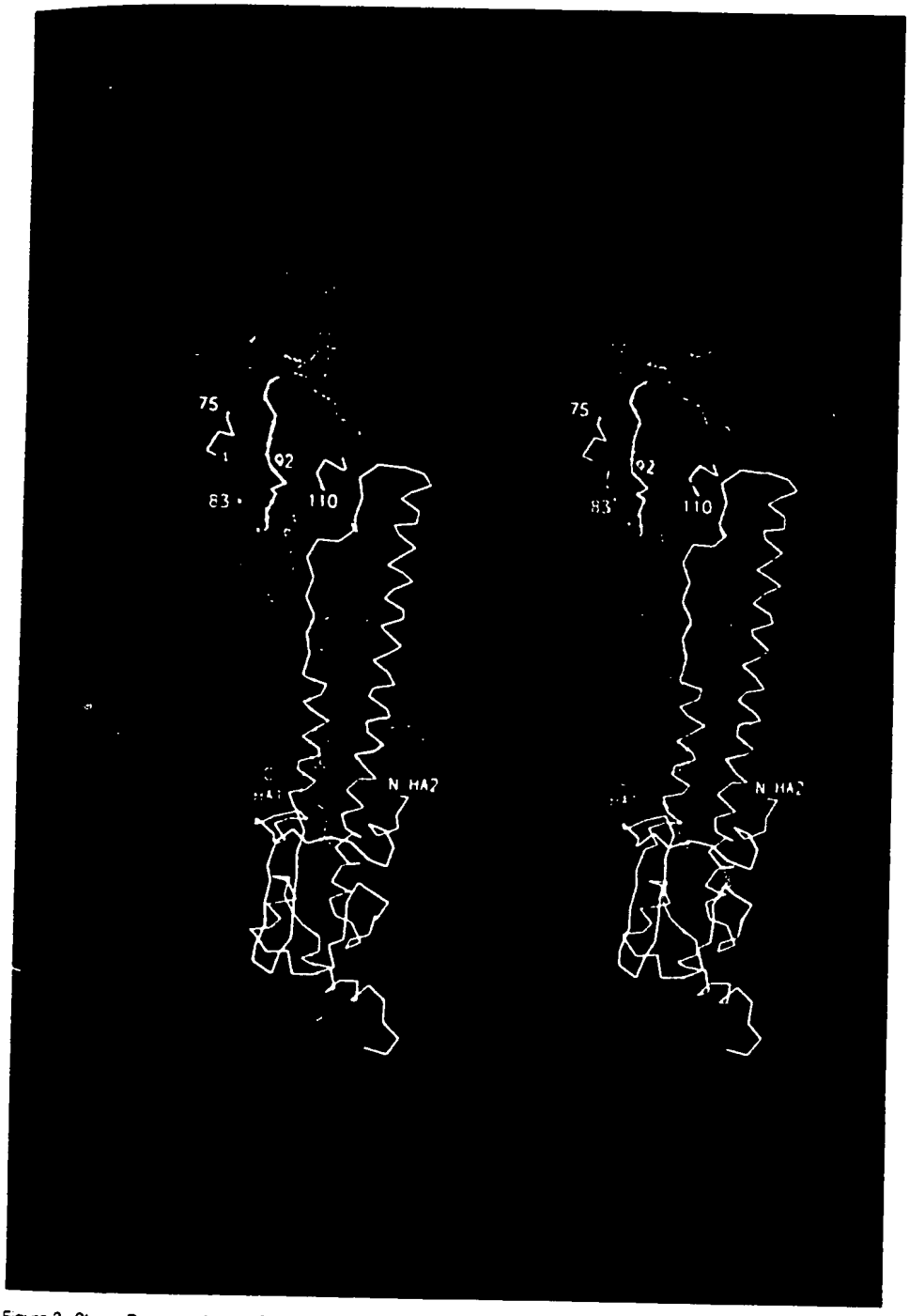

Figure 2. Stereo Drawing of the α-Carbon Trace of the X:31 Influenza Virus Hemagglutinin Monomer

The HA1 chain is shown in dark blue, and the HA2 chain is shown in light blue. The amino- and carboxyl-terminae of each chain is represented by N and C, respectively. The location of the sequence of the chemically synthesized peptide, residues 75–110, in the hemagglutinin molecule is shown in yellow, with sequences corresponding to sites 2 and 4 shown in red. This figure, as with Figures 3–7, was generated using the unpublished coordinates of Wilson, Skehel, and Wiley (see also Wilson et al., 1981).

surface area of 1293 sq. Å (to a probe of radius 3.0 Å) the peptide in the trimeric hemagglutinin molecule (Figure

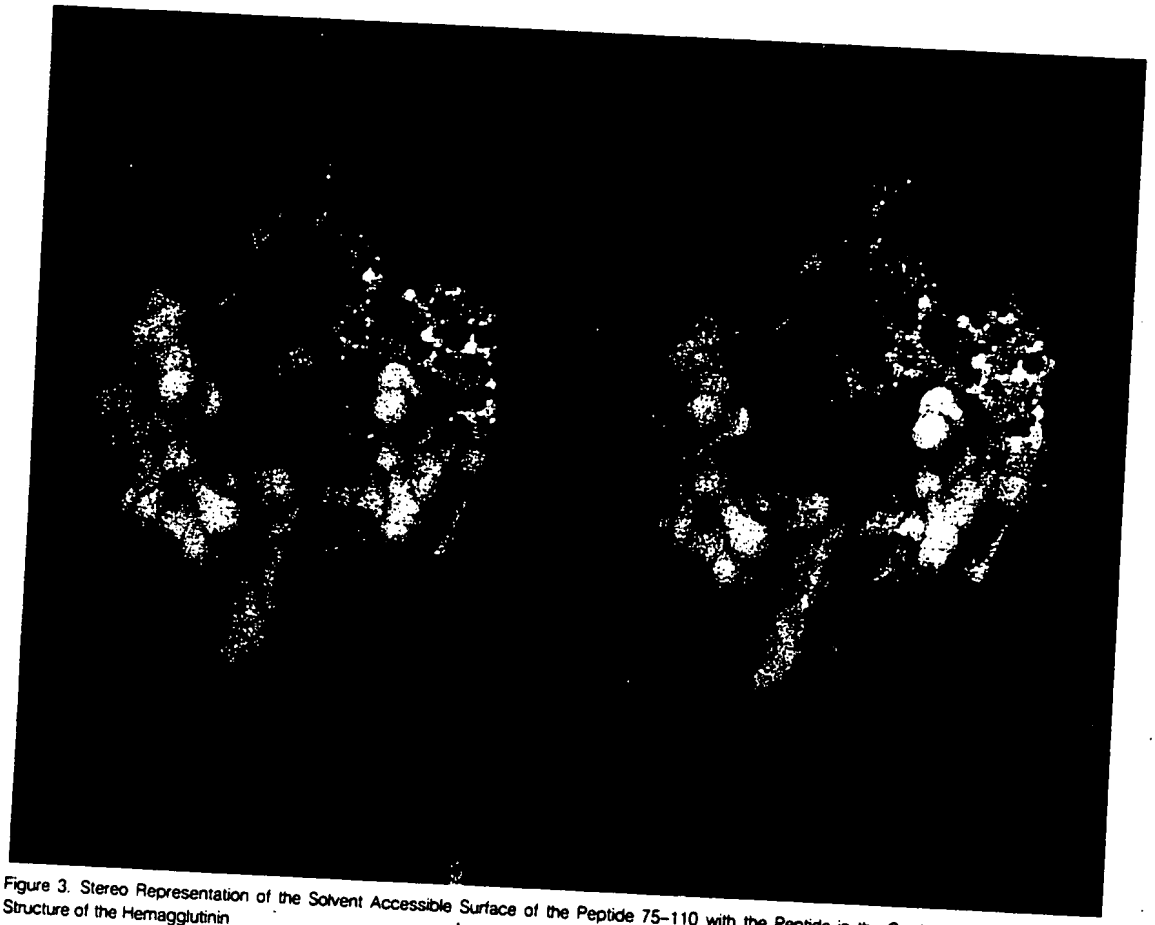

Figure 3. Stereo Representation of the Solvent Accessible Surface of the Peptide 75–110 with the Peptide in the Conformation It Adopts in the Crystal Structure of the Hemagglutinin

The conformation the peptide adopts in solution is unknown. Thus the figure represents a possible conformation of the peptide if the solution structure was similar to that in the folded protein. The peptide surface, as displayed on a raster display unit, is shown in yellow with the proposed immunodominant sequence, 98–106, shown by a gray translucent surface. The atoms of the peptide can be viewed through the translucent surface and are colored blue (nitrogen), red (oxygen), and green (carbon).

sents a continuous surface (Figure 5, blue surface), where residue 75 is close to residues 96 and 97, such that residues 96–97 and 75–83 form a continuous belt of accessible surface. The residues that are in the trimer interface (100–107, Figure 5, white surface) are accessible only to water molecules, small ligands, or molecules with small protruding surfaces. The sequence 83–91 has no accessible surface, as it is completely buried, even in the monomer.

Structure of the Antigenic Determinants in the Protein

Chemical identification of the residues involved in recognition of the synthetic peptide by the monoclonal antibodies together with analysis of the location and conformation of the peptide in the native structure can be combined to describe the structure of the antigenic determinants in the protein. Outlined in most detail are the peptide 98–106, as this is the sequence which appears to be immunodominant with the 75–110 peptide as immunogen, and the peptide 79–86, as a representative of the other three peptide sequences which are shown to be recognized by antibodies.

Site 1, Residues 75–86

The residues 75–83 are all mainly accessible on the surface with residues 76, 79, and 84–86 buried.

Site 2, Residues 79–86

The solvent-accessible atoms for the residues 79–86 are clustered around residues Gln 80, Asn 81, and Glu 82. The sidechains and mainchains of these residues form a shallow flat surface with dimensions 9 × 17 × 5Å (Figure 6). The other residues in this sequence are essentially buried except for the sidechain of Asp 85.

Site 3, Residues 83–92

This sequence is essentially buried except for some of the atoms of residues at each end. The data (Table 1) suggest binding is near the carboxyl end of this peptide. Arg 90, Ser 91, and Lys 92 form a sharp protrusion in the structure with approximate dimensions 14 × 9 × 8Å. If more than these three residues are recognized, the antibody would

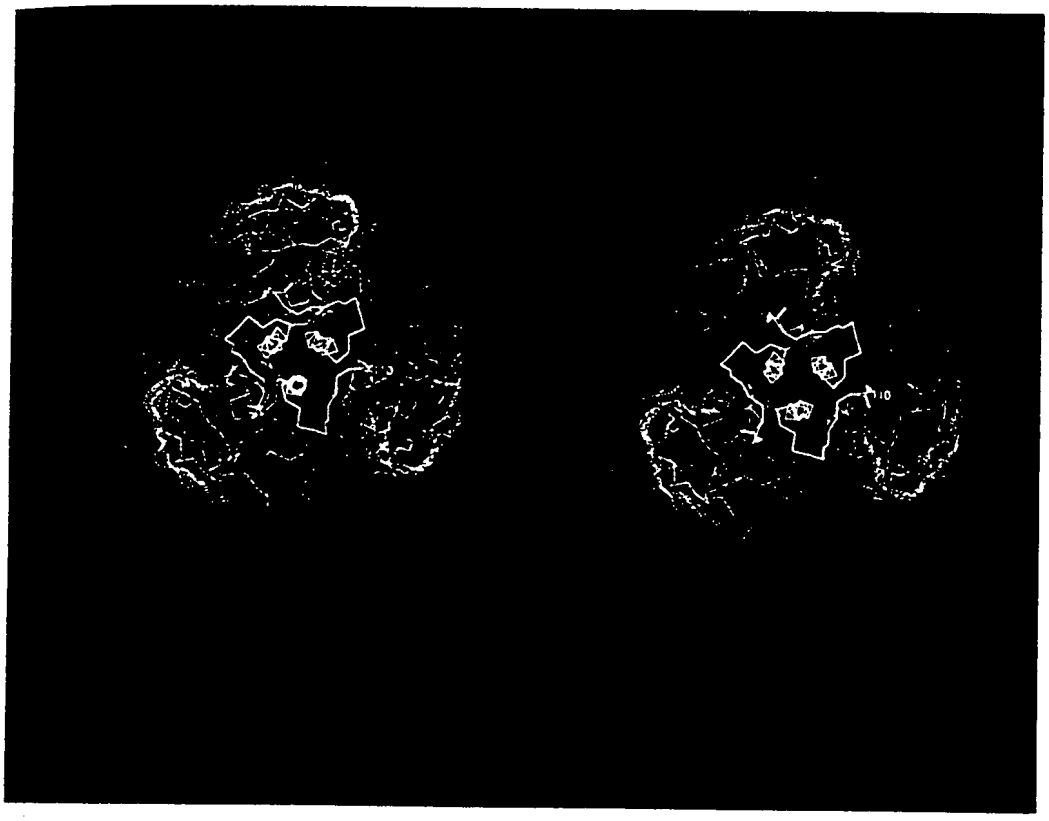

Figure 4. Stereo Drawing of an α-Carbon Trace of the Hemagglutinin Trimer with the Solvent-Accessible Surface of the Peptide 75–110

The globular heads of the trimer are shown looking down the 3-fold axis, from the end furthest from the membrane end of the hemagglutinin molecule. The peptide (75–110) α-carbon trace is shown in yellow with its solvent accessible surface represented by yellow dots, except for the immunodominant sequence (98–106) of the peptide, which is represented by red dots. The sequence 98–106 has residues accessible to a probe the size of a water molecule or small ions, but is inaccessible in this conformation to larger molecules such as immunoglobulins.

have to bind to a more unfolded structure of the protein molecule, as the other residues are inaccessible in the native conformation.

Site 4, Residues 98–106

This sequence lies in the trimer interface and has surface accessible only to small ligands for residues Tyr 100, Asp 101, Pro 103, Asp 104, Tyr 105, and Ala 106 (Figure 7). The peptide binding studies indicate that the binding can probably be localized to residues 100–106. These residues as with site 2 form a fairly flat, slightly convex surface, embedded in the rest of the protein surface, with approximate dimensions 16 × 15 × 7Å. This peptide contains two prolines, two aspartates, and three tyrosines and may be conformationally restricted as a result of the more limited torsional angles of proline residues. Unless the antibody hypervariable loops can protrude into the cavity in the trimer interface, the structure indicates that the sequence would be inaccessible in the trimer to antibody binding. Thus antibody binding is suggested to occur to a structure in which the hemagglutinin monomeric heads are exposed and which represents a conformation different from the native hemagglutinin trimeric structure

Discussion

The study of reactivity of overlapping peptide fragments with a panel of monoclonal antibodies has identified sequences within a 36 amino acid peptide that are immunodominant in producing hybridomas that react equivalently in ELISA binding studies with the intact hemagglutinin and virus. Thus an antigenic determinant, that is, the region of the molecule to which an antibody binds, can be identified not only for the synthetic peptide but also for the intact antigen. With caveats in mind about the uncertainty of the conformation of the free peptide (see Results), the question of which amino acids constitute an antigenic determinant and what is the nature of antibody–antigen union is advanced by these results.

One sequence of nine amino acids (98–106) is proposed to be immunodominant in producing antibodies against the 36 amino acid peptide. The findings agree with those of Muller et al. (1982) in that our sequence is contained in the 98–108 peptide used by them to generate antibodies against the A/Texas/1/77 (H3 subtype) strain of influenza. Only three of the panel of antibodies can be localized to

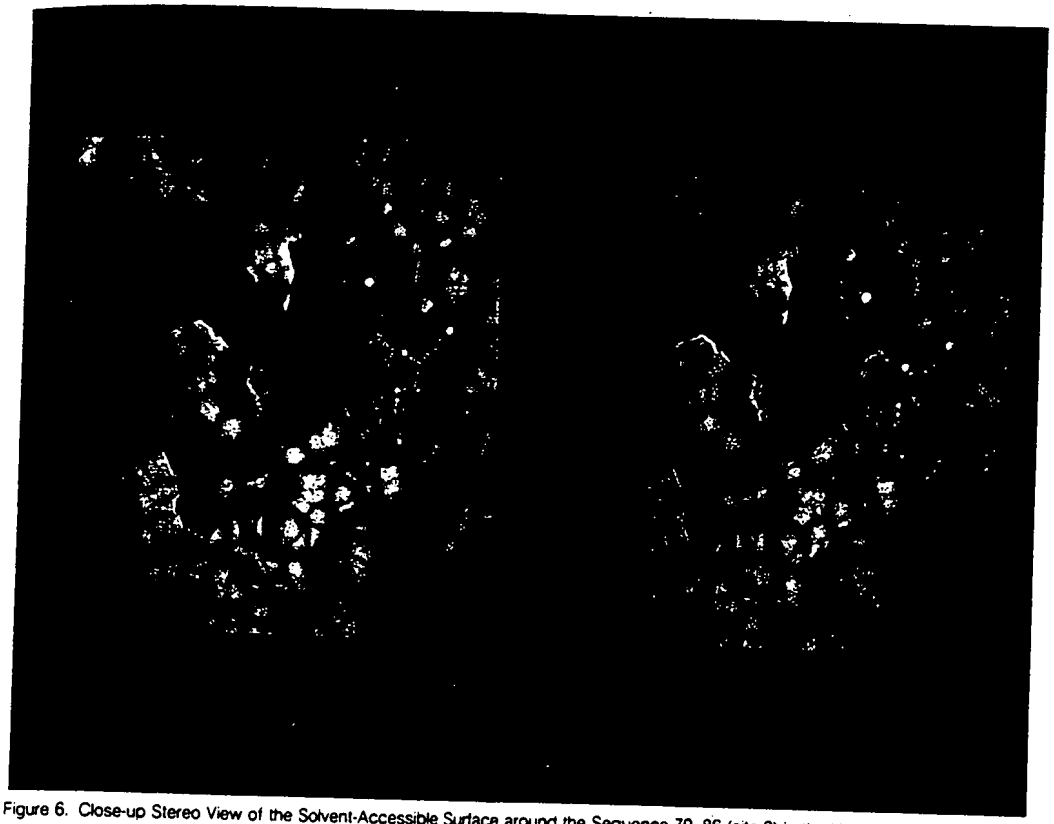

Figure 6. Close-up Stereo View of the Solvent-Accessible Surface around the Sequence 79–86 (site 2) in the Hemagglutinin Structure Displayed on a Raster Display Unit

A ball-and-stick representation of the solvent accessible atoms of the sequence 79–86 can be viewed through the translucent gray surface: The carbon, nitrogen, and oxygen atoms are colored green, light blue, and red, respectively. The non-translucent blue surface shows the other surrounding residues of the 75–110 sequence embedded in the HA1 polypeptide (yellow).

map to a different region (sites 1–3). Seven others react with the complete peptide, or larger fragments, and presumably are comprised of amino acids spatially separated on the sequence, as has been identified for antibody-combining sites of intact protein antigens such as the hemagglutinin (Wiley et al., 1981) and neuraminidase (Colman et al., 1983) of influenza virus.

The location of the antigenic sequences are on the surface, in the trimer interface, or essentially buried, even in the monomer. None of these sequences protrudes particularly from the surface in the native protein structure; they are presented as slightly convex surfaces comprised of 3–10 amino acids surrounded by the rest of the protein (Figures 6 and 7). These simple antigenic sites differ markedly from the complex nature of antibody-combining sites described after immunization with the intact hemagglutinin protein (Wiley et al., 1981) as well as other intact antigens (Crumpton, 1974; Benjamin et al., 1984).

It is interesting to consider the nature of antigenic determinants in terms of what is known about antigen-combining sites of immunoglobulin molecules. The structure determination of myeloma proteins (reviewed by Amzel and Poljak, 1979), either of intact IgG molecules or fragments such as Fab's, have shown the hypervariable regions to be located around a cleft at the distal ends of the molecule. Binding studies with myeloma proteins have shown that small ligands such as vitamin K (Amzel et al., 1974) or phosphorylcholine (Segal et al., 1974) bind to antibodies in this region but occupy only a small portion of the potential antigen-binding site. Characterization of the binding sites of hybridoma antibodies specific for $\alpha(1\rightarrow6)$ linked dextran showed that an antigen of about 5–7 hex-

Figure 5. The Solvent-Accessible Surface of a Hemagglutinin Trimer Represented on a Raster Display Unit Highlighting the Location of the Sequence 75–110 in the Folded Protein Structure

The left monomer shows HA1 in red, HA2 in gray, and the right monomer has HA1 in yellow and HA2 in light blue. The sequence 75–110 is represented by dark blue except for the subsequence 98–106, which is shown in white. The sequence clearly has residues accessible to the surface where the blue surface can be traced from the right monomers around the HA1/HA2 interface (yellow/light blue) to the back of the yellow monomer, shown on the adjacent subunit (red). The sequence 98–106 lies in the trimer interface and the relative lack of contacts of the trimeric heads allows this sequence to be visible although inaccessible to large ligands.

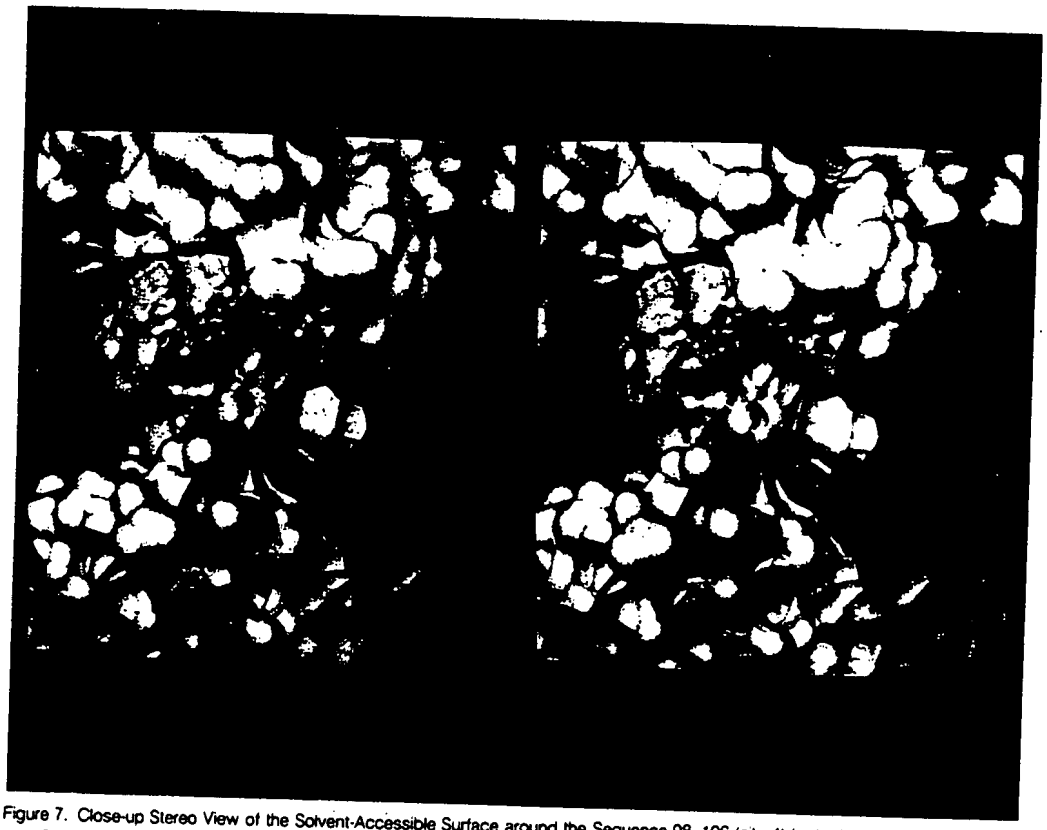

Figure 7. Close-up Stereo View of the Solvent-Accessible Surface around the Sequence 98–106 (site 4) in the Hemagglutinin Monomer Structure Displayed on a Raster Display Unit

The other two monomers of the trimer are removed for the surface calculation to enable this surface to be visible in close-up. A ball-and-stick representation of the solvent-accessible surface in the monomer of the sequence 98–106 can be viewed through the translucent gray surface. The C, N, and O atoms are colored green, light blue, and red, respectively. The non-translucent dark blue surface shows the other surrounding residues of the sequence 75–110 embedded the HA1 polypeptide chain (yellow) and close to the top of the long helix of HA2 (around residues 73–78) which form the start of the triple-coiled coil of the fibrous stem of the trimer.

oses was complementary to the binding site of IgA and IgG antibodies (Sharon et al., 1982). However, these anti-dextran antibodies have been shown to differ in their specificities in that some antibodies bind to the non-reducing end of the dextran whereas others bind to the middle six hexoses (Cisar et al., 1975; Kabat, 1978). These results have led to suggestions of different shapes of the antigen-combining sites as pockets (end-binding to dextran and McPc603 phosphorylcholine-binding antibodies) or lengthy grooves (six middle hexose-binding antibodies) or shallow grooves (Newm, menadione-binding antibodies) (see review by Davies and Metzger, 1983). Our calculation of the accessible surface of the pocket formed by the hypervariable residues of Fab Newm (Amzel et al., 1974) shows the potential binding surface to be 44 × 35 × 34Å in dimension with an area of 2400 sq. Å. The antigenic determinants of the free peptide reported here have dimensions in site 1 of 18 × 23 × 14Å, in site 2 of 13 × 22 × 12Å, in site 3 of 29 × 13 × 11Å, and in site 4 of 12 × 22 × 14Å and with accessible surface areas in the protein of 850–1160 sq. Å. Thus peptide determinants of these dimensions would occupy a sizeable portion of the antigen-binding site of the antibody molecule.

Finally, and perhaps most importantly, the ELISA binding data presented in this paper suggest that antibodies can bind to a structure different from that of the native trimeric antigen and consequently the protein may display conformational mobility at least in local regions. Conformational changes affecting the quaternary structure of the hemagglutinin molecule have been identified previously. For example, such changes have been implicated in the fusion activity of the virus with the target membrane of the host cell. Perturbation of the structure by low pH (pH 5.0–5.5. Skehel et al., 1982) results in a change in conformation such that the previously buried fusion peptide is extruded. The molecule can then be cleaved with trypsin to obtain intact monomeric heads that have selectively modified antigenic activity (Daniels et al., 1984). In addition, unless amino acid substitutions affect distant regions in the molecule, one of the proposed sites in the hemagglutinin to which neutralizing antibodies against the intact molecule bind (Site D, Wiley et al., 1981) would require binding to residues in the trimer interface and hence binding to a conformation different from that of the trimeric molecule Studies of apomyoglobin and holomyoglobin have also suggested that antibodies can bind to alternative confor-

mations of protein molecules (Crumpton, 1974). Upon reaction with antibodies, brown holomyoglobin formed a white precipitate, suggesting that antigen–antibody union shifted the equilibrium toward a protein structure that excluded the heme moiety.

A thorough understanding of how antipeptide antibodies bind to proteins will require further studies. The question of whether antibodies induce conformational changes (or "melt" the protein) or whether they bind to minor conformations which the protein has access to in solution has still to be answered. Clearly, successful antibody binding in ELISA on microtiter plates indicates that under these conditions both the peptide and protein can obtain conformations favorable for antibody–antigen union. To understand the detailed chemistry of antibody–antigen union in solution, we need to measure binding constants of the antibodies with each of the peptides as well as with the intact protein. Preliminary results with peptide competition studies already indicate that even the short peptides (≈9 amino acids) can bind and displace the larger peptides in immunoprecipitation assays, although with lower efficiency (Bergmann and Wilson, unpublished data).

The complete determination of the structure of peptide antigenic determinants awaits the determination not only of the free peptide conformation but also of the conformation of the peptide and of the intact protein when complexed with such an antipeptide antibody. Such studies, using x-ray crystallographic methods and N.M.R. spectroscopic techniques, are in progress here and in other laboratories.

Experimental Procedures

Synthesis of Peptides

Peptides were synthesized by the solid-phase method (Merrifield, 1964; Houghten et al., 1980), using a Beckman Model 990B peptide synthesizer. A representative synthesis follows. One gram (0.5 mEq approx.) of N-tert-butoxyl-carbonyl-amino-acid resin was used in conjunction with the following sidechain protecting groups: O-bromo-benzyloxycarbonyl(tyrosine and lysine), ε-N-tosyl (arginine), O-benzyl (glutamic acid, aspartic acid, threonine, and serine) and dinitrophenyl (histidine) and N-formyl (tryptophan). Protected amino acids were recrystallized from appropriate solvents and their purity verified by thin layer chromatography. The couplings were carried out with a 10-fold excess of protected amino acid and dicyclohexylcarbodiimide. An equimolar quantity of N-hydroxy-benzotriazole was added to the protected asparagine and glutamine BOC-amino acids, with dimethylformamide as solvent. All couplings were 99% complete as assayed by the picric acid test. The dinitrophenyl residue was removed from histidine with mercaptobenzene prior to hydrogen fluoride (HF) cleavage. The protected peptide polymers were then treated with twice their weight of anisole and 40 times their volume (relative to weight) of anhydrous HF for 1 hr at 4°C. After removal of the HF with nitrogen, the residue was extracted three times with ether (3 × 50 ml) and dried in vacuo. The resulting mixture of peptide and resin was extracted three times with 5% acetic acid, filtered, and dried in vacuo. The amino acid analysis indicated that all the peptides had amino acid compositions that were ±5% of their theoretical values.

The peptides were synthesized according to the amino acid sequence of X:47 (A/Victoria/3/75) influenza virus hemagglutinin predicted from the nucleotide sequence (Min Jou et al., 1980). In this paper, the sequence number corresponds to the predicted amino acid sequence of A/Aichi/2/68 (X 31) (Verhoeven et al., 1980) and the X 31 structure determination of

Preparation of Antipeptide Antibodies

A peptide (75–110) of HA1 hemagglutinin was used to immunize 129 GIX+ mice as described (Niman et al., 1983). Spleen cells were fused with SP2/0 myeloma cells with polyethylene glycol 1500 (Baker). The cells were resuspended with 400 ml of Dulbecco's high glucose minimal essential medium containing 10% fetal calf serum, 100 μM hypoxanthine, 1.0 μM methotrexate and 16 μM thymidine and plated into 30 microtiter plates and grown as described (Niman and Elder, 1980). The 18 cell lines produced are those previously described (Niman et al., 1983).

Antibody Binding Assays

Reactivity of the hybridomas was determined in enzyme-linked immunosorbent assays (ELISA). The antigen, diluted in phosphate-buffered saline (pH 7.4), was dried onto microtiter plates, fixed by methanol and incubated with tissue culture supernatant as described (Niman and Elder, 1982a). The microtiter plates were thoroughly washed and binding was detected by addition of rabbit anti-mouse κ-chain (Litton) and followed by a glucose oxidase-conjugated goat anti-rabbit antisera (Niman and Elder, 1982b).

Excess goat anti-rabbit antisera was washed off and the reaction was developed by adding ABTS dye (Boehringer, Mannheim) in the presence of glucose and horseradish peroxidase (Niman and Elder, 1982a). The plates were read in a Titertek Multiskan at A414 and the values presented as the amount of antigen at which 50% of the antibody was bound.

Correlation of Peptide Antigenic Sites with Hemagglutinin Structure

The x-ray structure on which the analysis of the location of peptides and their cognate structure in the intact Hong Kong 1968 hemaggutinin was as determined by Wilson et al. (1981). Previously described neutralizing antigenic determinants were as proposed by Wiley et al. (1981) and Daniels et al. (1984). The coordinates for the analysis were provided by Wilson, Skehel, and Wiley (unpublished data). Computer analysis of the peptide and protein structure was performed on a DEC VAX 11/750 and displayed on an Evans and Sutherland color Multi-Picture System. The structure representations were photographed directly from the Picture System using the GRAMPS (O'Donnell and Olson, 1981) and GRANNY (Connolly and Olson, submitted) and MS (Connolly, 1983a, 1983b) programs. The raster images were created by solvent accessible surface programs (Connolly, 1983a, 1983b) with recent enhancements (Connolly, unpublished) and displayed on an AED 767 frame-buffer raster display.

Acknowledgments

We are grateful to Diane Schloeder and Brian Heilman for excellent technical assistance. We thank Dr. John Skehel for providing purified bromelain-released hemagglutinin of X:47. I. A. W. acknowledges the valuable contribution of his colleagues Dr. Don C. Wiley of Harvard University and Dr. John J. Skehel of the National Institute for Medical Research, Mill Hill, London, for the unpublished x-ray coordinates used in this analysis. We thank Dr. James Hogle for helpful discussions, and Dan Haft and Ann McDonald for assistance in manuscript preparation. Support is acknowledged from National Institutes of Health grants AI 19499-01 (I. A. W. and R. A. L.) and CA 25803 (R. A. L.). This is publication number 3231-MB from the Research Institute of Scripps Clinic.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 3, 1984; revised March 9, 1984

References

Amzel, L. M., and Poljak, R. J. (1979). The three-dimensional structure of immunoglobulins. Ann. Rev Biochem. 48, 961–97

Amzel, L. M., Poljak, R. J., Saul, F., Varga, J. M., and Richards, F. (1974). The three-dimensional structure of a combining region–ligand complex of immunoglobulin New at 3.5Å resolution Proc. Nat. Acad. Sci. USA 71, 1427–1430

immunochemical anatomy of a protein and conclusions relating to antigenic structures of proteins. Immunochemistry *12*, 423–438.

Atassi, M. Z. (1978). Precise determination of the entire antigenic structure of lysozyme: molecular features of protein antigenic structures and potential of 'surface-stimulation synthesis'—a powerful new concept for protein binding sites. Immunochemistry *15*, 909–936.

Atassi, M. Z., Sakata, S., Kazin, A. L. (1979). Localization and verification by synthesis of five antigenic sites of bovine serum albumin. Biochem. J. *179*, 327–331.

Benjamin, D. C., Berzofsky, J. A., East, I. J., Gurd, F. R. N., Hannum, C., Leach, S. J., Margoliash, E., Michael, J. G., Miller, A., Prager, E. M., Reichlen, M., Sercarz, E. E., Smith-Gill, S. J., Todd, P. E., and Wilson, A. C. (1984). The antigenic structure of proteins: a reappraisal. Ann. Rev. Immunol. *2*, in press.

Berzofsky, J. A., Buckenmeyer, G. K., Hicks, G., Gurd, F. R. N., Feldmann, R. J., and Minna, J. (1982). Topographic antigenic determinants recognized by monoclonal antibodies to sperm whale myoglobin. J. Biol. Chem. *257*, 3189–3198.

Both, G. W., Sleigh, M. J., Cox, N. J., and Kendal, A. P. (1983). Antigenic drift in influenza virus H3 haemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes at key antigenic sites. J. Virol. *48*, 52–60.

Caton, A. J., Brownlee, G. G., Yewell, J. W., and Gerhard, W. (1982). The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell *31*, 417–427.

Cisar, J., Kabat, E. A., Dorner, M. M., and Liao, J. (1975). Binding properties of immunoglobulin combining sites specific for terminal or nonterminal antigenic determinants in dextran. J. Exp. Med. *142*, 435–459.

Colman, P. M., Varghese, J. N., and Laver, W. G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature *303*, 41–44.

Connolly, M. L. (1983a). Solvent-accessible surfaces of proteins and nucleic acids. Science *221*, 709–713.

Connolly, M. L. (1983b). Analytical molecular surface calculation. J. Appl. Crystallogr. *16*, 548–558.

Crumpton, M. J. (1974). Protein antigens: the molecular bases of antigenicity and immunogenicity. In The Antigens II, M. Sela, ed. (New York: Academic Press), pp. 1–78.

Daniels, R. S., Douglas, A. R., Gonsalves-Scarano, F., Palu, G., Skehel, J. J., Brown, E., Knossow, M., Wilson, I. A., and Wiley, D. C. (1984). Antigenic structure of influenza virus haemagglutinin. In The Origin of Pandemic Influenza Viruses, Chu and W. G. Laver, eds. (Amsterdam: Elsevier), in press.

Davies, D. R., and Metzger, H. (1983). Structural basis of antibody function. Ann. Rev. Immunol. *1*, 87–117.

East, I. J., Todd, P. E., and Leach, S. J. (1980). On topographic antigenic determinants in myoglobin. Mol. Immunol. *17*, 519–525.

Green, N., Alexander, H., Olson, A., Alexander, S., Shinnick, T. M., Sutcliffe, J. G., and Lerner, R. A. (1982). Immunogenic structure of the influenza virus hemagglutinin. Cell *28*, 477–487.

Houghten, R. A., Chang, W. C., and Li, C. H. (1980). Human B-endorphin: synthesis and characterization of iodinated and tritiated analogs. Int. J. Peptides Prot. Res. *16*, 311–320.

Hurrell, J. G. R., Smith, J. A., Todd, P. E., and Leach, S. J. (1978). Cross reactivity between mammalian myoglobins: linear vs. spatial antigenic determinants. Immunochemistry *14*, 283–288.

Jackson, D. C., Murray, J. M., White, D. O., Fagan, C. N., and Tregear, G. W. (1982). Antigenic activity of a synthetic peptide comprising the "loop" region of influenza virus haemagglutinin. Virology *120*, 273–276.

Kabat, E. A. (1978). The structural basis of antibody complementarity. Adv. Prot. Chem. *32*, 1–75.

Laver, W. G., Air, G. M., Webster, R. G., Gerhard, W., Ward, C. W., and Dopheide, T. A. (1979) Antigenic drift in type A influenza virus: sequence differences in the haemagglutinin of Hong Kong (H3N2) variants selected with monoclonal hybridoma antibodies. Virology *98*, 226–237

Laver, W. G., Air, G. M., Dopheide, T. A., and Ward, C. W. (1980). Amino acid sequence changes in the haemagglutinin of A/Hong Kong (H3N2) influenza virus during the period 1968–77. Nature *283*, 454–457.

Merrifield, R. B. (1964). Solid phase peptide synthesis. II. The synthesis of bradykinin. J. Am. Chem. Soc. *86*, 304–305.

Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Fang, R., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N., and Emtage, S. (1980). Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. Cell *19*, 683–696.

Muller, G. M., Shapira, M., and Arnon, R. (1982). Anti-influenza response achieved by immunization with a synthetic conjugate. Proc. Nat. Acad. Sci. USA *79*, 569–573.

Niman, H. L., and Elder, J. H. (1980). Molecular dissection of Rauscher virus gp70 by using monoclonal antibodies: localization of acquired sequences of related envelope gene recombinants. Proc. Nat. Acad. Sci. USA *77*, 4524–4528.

Niman, H. L., and Elder, J. H. (1982a). Structural analysis of Rauscher virus gp70 using monoclonal antibodies: sites of antigenicity and p15(E) linkage. Virology *123*, 187–205.

Niman, H. L., and Elder, J. H. (1982b). mAbs as probes of protein structure: molecular diversity among the envelope glycoproteins (gp70's) of the murine retroviruses. In Monoclonal Antibodies and T Cell Products, D. H. Katz, ed. (Boca Raton, Florida: CRC Press), pp. 21–51.

Niman, H. L., Houghten, R. A., Walker, L. A., Reisfeld, R. A., Wilson, I. A., Hogle, J. M., and Lerner, R. A. (1983). Generation of protein-reactive antibodies by short peptides is an event of high frequency: implications for the structural basis of immune recognition. Proc. Nat. Acad. Sci. USA *80*, 4949–4953.

O'Donnell, T. J., and Olson, A. J. (1981). GRAMPS—a graphics language interpreter for real-time interactive, three-dimensional picture editing and animation. Computer Graphics *15*, 133–142.

Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M., and Davies, D. R. (1974). The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. Proc. Nat. Acad. Sci. USA *71*, 4298–4302.

Sharon, J., Kabat, E. A., and Morrison, S. L. (1982). Immunochemical characterization of binding sites of hybridoma antibodies specific for $\alpha(1\rightarrow 6)$ linked dextran. Mol. Immunol. *19*, 375–388.

Skehel, J. J., Bayley, P. M., Brown, E. B., Martin, S. R., Waterfield, M. D., White, J. M., Wilson, I. A., and Wiley, D. C. (1982). Changes in the conformation of influenza virus haemagglutinin at the pH optimum of virus-mediated membrane fusion. Proc. Nat. Acad. Sci. USA *79*, 968–972.

Smith-Gill, S. J., Wilson, A. C., Potter, M., Prager, E. M., Feldmann, R. J., and Mainhart, C. R. (1982). Mapping the antigenic epitope for a monoclonal antibody against lysozyme. J. Immunol. *128*, 314–322.

Underwood, P. A. (1982). Mapping of antigenic changes in the haemagglutinin of Hong Kong influenza (H3N2) strains using a large panel of monoclonal antibodies. J. Gen. Virol. *62*, 153–169.

Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E., and Fiers, W. (1980). Antigenic drift between the haemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75. Nature *286*, 771–776.

Webster, R. G., and Laver, W. G. (1980). Determination of the number of nonoverlapping antigenic areas of Hong Kong (H3N2) influenza virus haemagglutinin with monoclonal antibodies and the selection of variants with potential epidemiological significance. Virology *104*, 139–148.

Wiley, D. C., Wilson, I. A., and Skehel, J. J. (1981). Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature *289*, 373–378.

Wiley, D. C., Wilson, I. A., and Skehel, J. J. (1984). The haemagglutinin membrane glycoprotein of influenza virus. In Biological Macromolecules and Assemblies, Volume 1, A. McPherson, ed. (New York: John Wiley), pp. 299–336.

Wilson, I. A., Wiley, D. C., and Skehel, J. J. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3Å resolution. Nature *289*, 366–373